



HUMAN GENOME EPIDEMIOLOGY (HuGE) REVIEW

Androgen Receptor CAG Repeats and Prostate Cancer

Kevin A. Nelson and John S. Witte

Prostate cancer is the most common nonskin malignancy and the second leading cause of cancer deaths among men in the United States. Prostate cancer ([Mendelian Inheritance in Man 176807]) has a complex etiology; presently, age, ethnicity, and family history are the most consistently reported risk factors associated with disease. Other potential risk and protective factors have also been suggested. Androgen, acting through the androgen receptor (AR) is helpful in preserving the normal function and structure of the prostate. The AR ([Mendelian Inheritance in Man 313700]) is a structurally conserved member of the nuclear receptor superfamily of ligand-activated transcription factors. Androgens, such as testosterone, are strong tumor promoters, and work with the AR to augment the effect of any carcinogens present and stimulate cell division. The CAG repeats encode long glutamine homopolymeric amino acid chains in the amino-terminal domain of the AR gene. The authors focus on CAG repeat length because recent research suggests that men with shorter AR CAG lengths (e.g., ≤ 22 repeats) are at a greater risk of developing prostate cancer than are those with longer variants. Among populations studied to date, African Americans appear to have the highest frequency of short CAG repeats. Several potential interactions have also been explored, including molecular interactions, androgen deprivation therapy, and prostate-specific antigen expression. CAG repeat length can be determined with high sensitivity and specificity. Presently, there is no recommended population screening for AR CAG repeat length. *Am J Epidemiol* 2002;155:883–90.

epidemiology; prostatic neoplasms; receptors, androgen; trinucleotide repeats

GENE

Androgens affect the human embryo in utero and lead to the development of male internal and external genitalia. During puberty, an increase in androgen levels leads to the start of spermatogenesis and growth of accessory sex organs such as the prostate gland (1). Androgen plays a role in prostate cancer growth. This is known because androgen is required in rodent induction models of prostate carcinogenesis because dogs and males castrated before puberty do not get prostate cancer and because androgen-ablative therapy inhibits prostate tumor growth during the time prior to the tumor reaching androgen independence (2). Unfortunately, within a short period of time, the tumor reaches a rapidly proliferating, hormone-independent state, resulting in adverse outcomes for the patient (3). It is not known exactly

how androgen is involved in prostate cancer etiology. If the androgen receptor (AR) had oncogenic potential, the androgen may play a role in initiation (4). Alternatively, androgen may be involved in promotion or progression (i.e., clonal expansion) by enhancing androgen-regulated processes, for example, growth and cellular activity (4).

The AR is made up of a C-terminal hormone-binding domain that helps with ligand specificity, a central DNA-binding domain that attaches to androgen-responsive target genes and an N-terminal domain that influences transcription efficiency (5). The AR gene is located at Xq11.2-q12 (markers DXS991-DXS983) and is more than 90 kb in length (6, 7). The open reading frame is separated over eight exons. The large amino-terminal domain is encoded by exon one, which includes the highly polymorphic CAG repeats (8). Exons two and three encode the DNA-binding domain, and exons four to eight encode the information for the steroid hormone- (ligand) binding domain (8, 9). Androgen, acting through the AR, helps to maintain the normal function and structure of the prostate (5, 10, 11). When androgens bind to the AR, it is activated, dimerizes, and localizes in the nucleus, where it attaches to specific sequences in the regulatory regions of target genes (12). Important functions of the product of the AR gene include activation of the expression of other genes (11) and

Received for publication October 4, 2000, and accepted for publication October 12, 2001.

Abbreviations: AR, androgen receptor; BPH, benign prostatic hypertrophy; CI, confidence interval; PCR, polymerase chain reaction; SBMA, spinal and bulbar muscular atrophy.

From the Department of Epidemiology and Biostatistics, Case Western Reserve University, 2109 Adelbert Road, Cleveland, OH 44106-4945 (e-mail: witte@darwin.cwru.edu). (Reprint requests to Dr. John S. Witte at this address).

the transport of the androgen hormone (13).

It has been observed that shorter *AR* CAG repeats impose a higher transactivation activity on the receptor and have an increased binding affinity for androgens (6, 14). This may make the prostate more vulnerable to chronic androgen overstimulation and increased proliferative activity, which, in turn, could increase the rate of somatic mutations among tumor suppressor genes (e.g., the *AR*). In contrast, the expansion of the CAG repeat (>40 repeats) leads to a below-normal *AR* level when measured by transient transfections assays with hormone-responsive-computerized axial tomography constructs (15–18) and by Scatchard analysis of *AR* in skin fibroblasts (19). It has also been found that there is altered coactivator interaction with the *AR* (3). Thus, the CAG microsatellite encodes a polyglutamine tract, which has a length that is inversely and linearly related to *AR* activity (20, 21).

GENE VARIANTS

Table 1 presents the frequency by ethnicity of *AR* CAG repeat lengths among relevant studies detected in a search of Medline between 1990 and September 2001. We linked the keywords “androgen receptor” and “trinucleotide repeats” for one search and the keywords “androgen receptor” and “prostatic neoplasms” for another. (This search strategy was also applied to the remaining sections of this paper.) Note that estimates of population-specific allele frequencies might be affected by misclassification of ethnicity. For example, Edwards et al. (22) give estimates of CAG repeat frequency among European Americans, African Americans, Asian Americans, and Latinos using DNA extracted from a convenience sample of blood donated to Houston, Texas, blood banks. Blood-bank personnel noted European-American and African-American ethnicity based on visual appearances. Asian-American and Latino ethnicities were approximated based on surnames. Using visual or name identification could have led to severe misclassification of ethnicity, resulting in under- or overestimates of CAG repeat frequencies.

There is some evidence that CAG repeats are in linkage disequilibrium with other polymorphisms in the *AR* locus. CAG repeats were found to be in disequilibrium (i.e., allelic association) with Stu I mutations, and both have been associated with prostate cancer (13). The disequilibrium may, however, simply reflect conserved haplotypes arising from the close physical distance between these polymorphisms, since they are both in the *AR* gene. Irving et al. (24) also found evidence of linkage disequilibrium between CAG and GGC/N alleles among prostate cancer patients. However, Stanford et al. (11) were unable to demonstrate such an association and speculate that a nearby gene may be in linkage disequilibrium with the *AR*. Finally, Xue et al. (25) speculated that short CAG alleles may be in linkage disequilibrium with the prostate-specific antigen G allele.

DISEASES

Incidence and mortality

Prostate cancer is the most common solid tumor and the second leading cause of cancer deaths among men in the

United States: In 2001, approximately 198,100 men were diagnosed with prostate cancer, and about 31,500 will die of the disease. On the basis of data gathered during 1994–1996, a male has a one-in-six chance of developing prostate cancer at some point in his lifetime (26).

Ethnicity

Ethnicity is also one of the most widely accepted risk factors for prostate cancer. This helps to explain, in part, the wide geographic variation in the incidence of this disease. For instance, the United States has an incidence of prostate cancer that is eight times higher than that of Japan (27), and African Americans have the highest incidence and mortality in the world. For example, the Surveillance, Epidemiology, and End Results program data from the years 1988–1992 estimate age-adjusted incidence rates of 181/100,000 for African-American men versus only 135/100,000 among Caucasians; similarly, the age-adjusted prostate cancer mortality rate for African Americans is 54/100,000 versus only 24/100,000 in Caucasians (28). In all likelihood, racial and geographic differences are multifactorial in nature, with environmental, genetic, and possibly social factors (29, 30). Differences in diet and access to improved detection methods could explain some of the ethnic variation (31). Some have also suggested that varying levels of testosterone and its intraprostatic metabolite dihydrotestosterone across ethnic groups may be responsible for these differences (32).

Family history

In addition, family history is a generally accepted risk factor for prostate cancer. For younger men, familial factors appear to be particularly influential, and the attributable risk of strong familial factors may be as high as 43 percent for men who are less than age 55 years (33). Nevertheless, family history of prostate cancer accounts for only about 9 percent of all cases (33). First-degree relatives of men diagnosed with prostate cancer have a two- to sixfold excess risk of the disease (34–38). Spitz et al. (37) and Steinberg et al. (39) suggest that the more closely a man is related to an affected relative and the greater the number of affected relatives in his family, the higher is his risk of acquiring prostate cancer. Narod et al. (40) and Monroe et al. (41) found that brothers of cases faced significantly higher risks of prostate cancer than did fathers.

Other risk factors

Age is an important risk factor for prostate cancer (30): Men are rarely diagnosed before age 40 years, and the rate of increase with aging—at approximately the ninth to tenth power of age—is greater than that for any other cancer (13). A large number of other putative risk factors for prostate cancer have been identified, although whether these are causal remains controversial. Among these are obesity (42), alcohol (43), sexual behavior (35), cigarette smoking (44), use of pesticides (45), and nuclear power (46). Other factors (for example, selenium (47); vitamins A, D, and E (48–51);

and phytoestrogens (52, 53)) have been inversely associated with prostate cancer.

AR and prostate cancer

Prostate cancer is one of the primary diseases that have been associated with somatic and germline polymorphisms in the *AR* gene. In the "Associations" section below, we discuss in detail the relation between CAG repeats in the *AR* gene and prostate cancer. Other *AR* polymorphisms potentially associated with prostate cancer include GGC/N repeats, Stu I, and numerous base substitutions (6, 11, 13, 24, 50, 54–56).

CAG polymorphisms in the AR

Shorter CAG repeats have been associated with benign prostatic hyperplasia (BPH) (57, 58). Short CAG repeat lengths have also been associated with androgenetic alopecia (59, 60), ankylosing spondylitis (61), mental retardation (62), younger age of male rheumatoid arthritis (63), and hepatitis B-related hepatocellular carcinoma (64). In contrast, extremely long *AR* CAG length (i.e., 38–52 repeats) has been associated with Kennedy's disease, spinal and bulbar muscular atrophy (SBMA) (65, 66). SBMA is an adult-onset, motor neuron disease that is associated with reduced fertility, low virilization, testicular atrophy, and reduced sperm production (13). In addition, it is of interest to note that men with SBMA have also been diagnosed with prostate cancer (67). Expansions of the CAG repeat may also be involved in the carcinogenesis of uterine endometrial cells (68). Finally, Rebbeck et al. (69) found that *BRCA1* mutation carriers who carry long CAG repeats have an earlier age of breast cancer diagnosis. However, Spurdle et al. (70) found no association between CAG repeat length and breast cancer risk in women under age 40 years.

ASSOCIATIONS

CAG repeat length and occurrence of prostate cancer

Table 1 also gives odds ratios for the association between CAG repeats in the *AR* gene and prostate cancer. These results suggest that the length of CAG repeats might be inversely associated with the risk of prostate cancer. Coetzee and Ross (14) initially suggested that variations in CAG repeat length are associated with prostate cancer. One possible explanation for this is that short alleles may impose higher transactivation on the receptor due to the inverse relation between the number of glutamine residues in the polyglutamine tract and transcriptional activity (6, 14, 18). This potential association may be modified by stage/grade at diagnosis (71), and men diagnosed at an older age appear to have longer CAG repeats (9), suggesting that *AR* CAG repeat length may also be associated with prostate cancer aggressiveness. The use of a lower cutpoint shifts men from the short to the long category of trinucleotide repeat length. Thus, the proportion of short alleles is reduced, while that of longer alleles is increased. Other things being equal, this would bias the association between shorter alleles and the occurrence of prostate cancer. One would also

expect that the use of a lower cutpoint would accentuate short-long allele differences between racial groups because a higher proportion of African Americans have a low number of CAG repeats (9–17) compared with Caucasians or Asians. Reducing the short-long cutpoint should reclassify a greater proportion of Caucasians from the short to the long group compared with African Americans. The ratio of proportions for the shorter alleles for African Americans compared with Caucasians is higher in the study by Sartor et al. (72) than it is for the other studies that would be expected since a lower cutpoint is used.

Two studies listed here (73, 74) have not detected a statistically significant association between shorter CAG length and the occurrence of prostate cancer. Bratt et al. (75) concluded that CAG repeats are associated with young age at diagnosis, but not with higher risk of the disease. This study did not give the distribution of CAG repeat lengths among the controls in tabular form. As a result, the odds ratio could not be computed, and this study was not included in table 1. Eeles et al. (76) did not find an association between CAG repeat length and prostate cancer development or aggressiveness among 178 Caucasian cases and matched controls. (Note that this has been published only in abstract form and does not give sufficient details for inclusion in table 1).

Most studies reported to date suggest a positive association between shorter CAG repeat lengths and prostate cancer. However, the magnitude of any potential association appears to be relatively limited, and small sample sizes have restricted the precision of published results. Nevertheless, the large population frequency of short CAG repeat lengths suggests that, if causal, variants in CAG lengths could have substantial public health implications. More specifically, this potentially low-penetrance, high-frequency mutation could theoretically account for many more prostate cancer cases than a high-penetrance, low-frequency mutation.

CAG repeat length and aggressiveness of prostate cancer

With regard to prostate cancer aggressiveness, Giovannucci et al. (71) found that comparing CAG repeat lengths among subjects with a high grade/stage ($n = 269$) gave an odds ratio of 1.64 (95 percent confidence interval (CI): 1.22, 2.22), whereas those with a low grade/stage ($n = 309$) had an odds ratio of 1.02 (95 percent CI: 0.77, 1.37). Stanford et al. (11) also looked at the impact of CAG repeat length on disease aggressiveness (i.e., stage C or D or Gleason score 8–10). Comparison of short versus long CAG repeats (cutpoint = 22) gave an odds ratio of 1.20 (95 percent CI: 0.79, 1.84) for more aggressive prostate cancer. These results suggest that shorter CAG repeat length may be involved not only with the development of prostate cancer but also the potential aggressiveness of the disease.

CAG repeat length and BPH

A recent study by Giovannucci et al. (57) found that shorter CAG repeats are associated with BPH. The data for this came from the Health Professionals Follow-up Study,

TABLE 1. Androgen receptor CAG repeat length and association with prostate cancer

Study (reference)	Place of study	Description of controls	Description of cases	Ethnic group	No. of subjects	CAG repeat frequency		CAG repeat outpoint	Cases		Controls		OR*	95% CI*
						Short	Long		No.	Frequency	No.	Frequency		
Edwards et al. (22)	Houston, TX			Caucasian	240	0.62	0.38	>22						
				African American	250	0.76	0.24	≤22						
				Asian	97	0.55	0.45							
				Hispanic	205	0.48	0.52							
Irving et al.† (24)	Los Angeles, CA	Cross-sectional survey	Prostate cancer patients identified by SEER*	Caucasian	39	0.63	0.38	≥22	19	0.33	15	0.38	1.00	Referent 0.54, 2.92
				African American	44	0.75	0.25	<22	38	0.67	24	0.62	1.25	
Ingles et al.† (50)	Los Angeles, CA	Population-based case-control study of bladder cancer, matched to cases on age, sex, and race	Same as Irving et al. (24)	Asian	39	0.49	0.51	≥22	19	0.33	68	0.40	1.00	Referent 0.72, 2.54
				Non-Hispanic White	169	0.60	0.40	<22	38	0.67	101	0.60	1.35	
Giovannucci et al. (71)	US	From the Physician's Health Study	From the Physician's Health Study	Caucasian (over 95%)	588	0.48	0.52	≥22 <22	274 313	0.47 0.53	306 282	0.52 0.48	1.00 1.24	Referent 0.99, 1.56
Stanford et al. (11)	King County, WA	Population-based; ascertained by age-stratified random-digit telephone dialing	Population-based	Caucasian	266	0.48	0.52	≥22	136	0.48	140	0.53	1.00	Referent 0.84, 1.65
								<22	145	0.52	126	0.47	1.18	

Hakimi et al. (23)	US	From three "samplings" in different geographical areas	Men undergoing radical retropubic prostatectomy	Caucasian	370	0.03	0.97	>17 ≤17	53 6	0.90 0.10	359 11	0.97 0.03	1.00 3.7	Referent 1.3, 10.4
Sartor et al. (72)	US	Group of unaffected men		Caucasian	130 65	0.28 0.57	0.72 0.43	>20 <20						
Edwards et al. (74)	London, England	Women from a previous study of breast cancer	Referred cases systematically ascertained	Caucasian	390	0.54	0.46	≥22 <22	74 88	0.46 0.54	178 212	0.46 0.54	1.00 1.00	Referent 0.96, 1.03
Correa-Cerro et al. (73)	Paris and Nancy, France	Controls free of prostate cancer over 7 years	Consecutive series	French-German	105	0.52	0.48	≥22 <22	69 63	0.52 0.48	50 55	0.48 0.52	1.00 0.83	Referent 0.73, 0.95
Ekman et al. (30)	Stockholm, Sweden; Yamaguchi, Japan	Men with BPH*	From families with clustering of PCa	Swedish Japanese	38 33	0.24 0.24	0.76 0.76	≥15 <15	44 15	0.75 0.25	29 9	0.76 0.24	1.00 1.1	Referent 0.43, 2.84
Lange et al. (84)	MI (cases) Los Angeles, CA and King County, WA (controls)	Same as used by Irving et al. (24) and Stanford et al. (11)	From families with two or more cases of prostate cancer	Caucasian	133			>21 ≤21					1.00 0.99	Referent 0.60, 1.36

* OR, odds ratio; CI, confidence interval; SEER, Surveillance, Epidemiology, and End Results; BPH, benign prostate hypertrophy.

† These two studies used the same cases, with the paper by Irving et al. (24) simply adding more controls. In the paper by Edwards et al. (22), DNA was taken from blood banks from unrelated volunteers in a multilocus genotype survey. In the paper by Sartor et al. (72), subjects who had signs and symptoms of prostate cancer were excluded from the study. In the paper by Hakimi et al. (23), there was no listing of the number of observations per measured repeat length, which is necessary for recategorizing the data at a higher cutpoint (e.g., 22 CAG repeats), such as that used in other studies (e.g., the paper by Ingles et al. 50). In the paper by Lange et al. (84), the individual allele data were not provided, so we could not complete all columns.

an ongoing nationwide prospective cohort study of men aged 40–75 years. The subjects were first enrolled in 1986 and were employed in a variety of health professions. This study used 349 cases and 449 controls, and an odds ratio of 1.92 (95 percent CI: 1.22, 3.03) for BPH was observed for subjects with short alleles (<19 repeats) compared with those with long alleles (≥ 25 repeats) (57).

The potential public health relevance and the biologic rationale supporting the relation between *AR* CAG and prostate cancer has helped generate a broad interest in CAG repeats. A recent paper in *Epidemiologic Reviews* provides additional information about CAG repeats and prostate cancer (77). When results from studies that are investigating the potential associations between trinucleotide repeats and prostate cancer are available, we will incorporate them into the electronic version of this Human Genome Epidemiology review. (Results can be sent directly to Dr. John Witte at witte@darwin.cwru.edu.) Since numerous different cutpoints have been used to distinguish between short and long *AR* CAG and GGC/N repeats, we suggest that future reports give crude data that allow for using a range of repeat cutpoints.

INTERACTIONS

Epidemiologic interactions

Some speculate that there could be an interaction between genetic factors and the ratio of estrogen to testosterone. This ratio increases with age and could affect the up-regulation of *AR* and, ultimately, the slope of age-incidence curves (13). Moreover, androgen deprivation therapy, a common therapy used in the later stages of prostate cancer, has been associated with the *AR*. In particular, amplification of the *AR* gene may come about during therapy and encourage tumor cell growth in a low-androgen-concentration environment (78).

Molecular and genetic interactions

Inappropriate RNA-binding protein interaction with specific messenger RNA may disturb cellular dynamics or change the regulation, transport, and expression of CAG containing RNA. This, in turn, may account for the observation of an inverse correlation between CAG repeat length and *AR* messenger RNA expression in studies involving lower primates. Furthermore, the CAG polyglutamine region may interact with some cellular proteins, and differences in the polyglutamine length could affect the tenacity of such interactions (79). There may also be interaction between CAG and GGN repeats. For example, when the subgroup in which both (CAG) n and (GGN) n alleles were short (CAG, <22; GGN, ≤ 16) was compared with those in which both alleles were long (CAG, ≥ 22 ; GGN, >16), an odds ratio of 2.05 (95 percent CI: 1.09, 3.84) was observed (11). Finally, there is evidence that allelic differences in *AR*-driven, prostate-specific antigen expression may affect prostate cancer risk (25).

The study of interactions between CAG repeats and environmental and/or genetic mechanisms in prostate cancer is hampered by lack of power. These limitations have been highlighted by Smith and Day (80) and observed in another the Human Genome Epidemiology review by Cotton et al. (81).

Laboratory tests

To determine *AR* CAG repeat length, extracted undigested DNA is amplified by polymerase chain reaction (PCR) in a series of two rounds using nested primers surrounding the repeat in exon 1 (24). The final products can then be analyzed by electrophoresis on a 5 percent denaturing polyacrylamide gel and subjected to autoradiography. One can then obtain the number of CAG repeats from the size of the predominant PCR product relative to a series of previously determined CAG size standards. Subsequently, all samples are ranked by size and reanalyzed by electrophoresis and autoradiography, with each allele of an equivalent size next to each other to validate the original size estimates (24). Bharaj et al. (82) improved on this technique somewhat by utilizing a fully automated system for the electrophoretic separation of the PCR product, utilizing fluorescent (nonisotopic) detection of standard-sized markers and using fragment analysis to establish the length of the PCR product. Estimates of sensitivity and specificity for this laboratory test have not been reported. Nevertheless, since the primers for this test are designed specifically for the CAG repeat, the chance of making an error in CAG number counts is very low, and the sensitivity and specificity are extremely high; the main source of error is likely to come from human origins, for example, incorrectly mixing or labeling samples (S. Ingles, personal communication, 1999). Wada et al. (83) describe a promising, but not widely used, alternative to the above method. This involves amplification via PCR, using a simple purification procedure and counting CAG repeats via matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.

Population testing

On the basis of the evidence summarized here, testing for CAG trinucleotide repeat length in the general population as part of a population screening program is not presently warranted. This underscores the need for doing additional research on this topic. The association between *AR* CAG repeat length and prostate cancer is intriguing but is still somewhat limited. To understand prostate carcinogenesis better, we should work to fully explicate any relation among genes such as CAG repeats and environmental factors that may interact to increase prostate cancer risk. The resulting information may ultimately improve our ability to prevent the development of prostate cancer and provide us with better prognostic information following diagnosis.

ACKNOWLEDGMENTS

Supported by grants from the National Institutes of Health (CA88164), DOD Prostate Cancer Research Program (DAMD17-98-1-8589), and the Urologic Research Foundation.

REFERENCES

1. Yong EL, Lim J, Qi W, et al. Molecular basis of androgen receptor diseases. *Ann Med* 2000;32:15–22.
2. Wilding G. The importance of steroid hormones in prostate cancer. *Cancer Surv* 1992;14:113–30.
3. Lamb DJ, Weigel NL, Marcelli M. Androgen receptors and their biology. *Vitam Horm* 2001;62:199–230.
4. Hakimi JM, Rondinelli RH, Schoenberg MP, et al. Androgen-receptor gene structure and function in prostate cancer. *World J Urol* 1996;14:329–37.
5. Barrack ER. Androgen receptor mutations in prostate cancer. *Mt Sinai J Med* 1996;63:403–12.
6. Feldman D. Androgen and vitamin D receptor gene polymorphisms: the long and short of prostate cancer risk. *J Natl Cancer Inst* 1997;85:1571–9.
7. Cussenot O, Valeri A, Berthon P, et al. Hereditary prostate cancer and other genetic predispositions to prostate cancer. *Urol Int* 1998;60 (suppl 2):30–4.
8. Trapman J, Cleutjens KB. Androgen-regulated gene expression in prostate cancer. *Semin Cancer Biol* 1997;8:29–36.
9. Hardy DO, Scher HI, Bogenreider T, et al. Androgen receptor CAG repeat lengths in prostate cancer: correlation with age of onset. *J Clin Endocrinol Metab* 1996;81:4400–5.
10. Dorkin TJ, Neal DE. Basic science aspects of prostate cancer. *Semin Cancer Biol* 1997;8:21–7.
11. Stanford JL, Just JJ, Gibbs M, et al. Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res* 1997;57:1194–8.
12. Chang C, Saltzman A, Yeh S, et al. Androgen receptor: an overview. *Crit Rev Eukaryot Gene Expr* 1995;5:97–125.
13. Ross RK, Pike MC, Coetzee GA, et al. Androgen metabolism and prostate cancer: establishing a model of genetic susceptibility. *Cancer Res* 1998;58:4497–4504.
14. Coetzee GA, Ross RK. Re: "Prostate cancer and the androgen receptor." (Letter). *J Natl Cancer Inst* 1994;86:872–3.
15. Mhatre AN, Trifiro MA, Kaufman M, et al. Reduced transcriptional regulatory competence of the androgen receptor in X-linked spinal and bulbar muscular atrophy. *Nat Genet* 1993;5:184–8.
16. Shimada N, Sobue G, Doyu M, et al. X-linked recessive bulbo-spinal neuronopathy: clinical phenotypes and CAG repeat size in androgen receptor gene. *Muscle Nerve* 1995;18:1378–84.
17. Nakajima H, Kimura F, Nakagawa T. Effects of androgen receptor polyglutamine tract expansion on proliferation of NG108-15 cells. *Neurosci Lett* 1997;222:83–6.
18. Tut T, Ghadeesy F, Trifiro M, et al. Long polyglutamine tracts in the androgen receptor are associated with reduced *trans*-activation, impaired sperm production, and male infertility. *J Clin Endocrinol Metab* 1997;82:3777–82.
19. MacLean HE, Warne GL, Zajac JD. Spinal and bulbar muscular atrophy: androgen receptor dysfunction caused by a trinucleotide repeat expansion. *J Neurol Sci* 1996;135:149–57.
20. Chamberlain NC, Driver ED, Mielsfeld RL. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucl Acid Res* 1994;22:3181–6.
21. Kazemi-Esfarjani P, Trifiro MA, Pinsky L. Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: possible pathogenic relevance for the (CAG)*n*-expanded neuropathies. *Hum Mol Genet* 1995;4:523–7.
22. Edwards A, Hammond HA, Jin L, et al. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 1992;12:241–53.
23. Hakimi JM, Schoenberg MP, Rondinelli RH, et al. Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clin Cancer Res* 1997;3:1599–1608.
24. Irving RA, Yu MC, Ross RK, et al. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Res* 1995;55:1937–40.
25. Xue W, Irvine RA, Yu MC, et al. Susceptibility to prostate cancer: interaction between genotypes at the androgen receptor and prostate-specific antigen loci. *Cancer Res* 2000;60:839–41.
26. Greenlee RT, Murray T, Bolden S, et al. Cancer statistics, 2000. *CA Cancer J Clin* 2000;50:7–33.
27. Takahashi H, Furusato M, Allsbrook WC Jr, et al. Prevalence of androgen receptor gene mutations in latent carcinomas from Japanese men. *Cancer Res* 1995;55:1621–4.
28. Parker SL, Davis KJ, Wingo PA, et al. Cancer statistics by race and ethnicity. *CA Cancer J Clin* 1998;48:31–48.
29. Haas GP, Sakr WA. Epidemiology of prostate cancer. *CA Cancer J Clin* 1997;47:273–87.
30. Ekman P, Gronberg H, Matsuyama H, et al. Links between genetic and environmental factors and prostate cancer risk. *Prostate* 1999;39:262–8.
31. Dijkman GA, Debruyne FM. Epidemiology of prostate cancer. *Eur Urol* 1996;30:281–95.
32. Latil A, Lidereau R. Genetic aspects of prostate cancer. *Virchows Arch* 1998;432:389–406.
33. Giovannucci E. How is individual risk for prostate cancer assessed? *Hematol Oncol Clin North Am* 1996;10:537–48.
34. Woolf CM. An investigation of the familial aspects of carcinoma of the prostate. *Cancer* 1960;13:739–44.
35. Steele R, Lees RE, Kraus AS, et al. Sexual factors in the epidemiology of cancer of the prostate. *J Chronic Dis* 1971;24:29–37.
36. Fincham SM, Hill GB, Hanson J, et al. Epidemiology of prostate cancer: a case-control study. *Prostate* 1990;17:189–206.
37. Spitz MA, Currier RD, Fueger JJ, et al. Familial patterns of prostate cancer: a case-control analysis. *J Urol* 1991;146:1305–7.
38. Goldgar DE, Easton DF, Cannon-Albright LA, et al. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst* 1994;86:1600–8.
39. Steinberg GD, Carter BS, Beaty TH, et al. Family history and the risk of prostate cancer. *Prostate* 1990;17:337–47.
40. Narod SA, Dupont A, Cusan L, et al. The impact of family history on early detection of prostate cancer. (Letter). *Nat Med* 1995;1:99–101.
41. Monroe KR, Yu MC, Kolonel LN, et al. Evidence of an X-linked or recessive genetic component to prostate cancer risk. *Nat Med* 1995;1:827–9.
42. Gronberg H, Damber L, Damber JE. Total food consumption and body mass index in relation to prostate cancer risk: a case-control study in Sweden with prospectively collected exposure data. *J Urol* 1996;155:969–74.
43. International Agency for Research on Cancer. In: Parkin DM, Muir GS, Whelan SL, et al., eds. Cancer incidence in five continents. IARC scientific publication no.120. Lyon, France: International Agency for Research on Cancer, 1992.
44. Hsing AW, McLaughlin JK, Hrubec Z, et al. Tobacco use and prostate cancer: 26-year follow-up of US veterans. *Am J Epidemiol* 1991;133:437–41.
45. Wiklund K, Dich J, Holm LE, et al. Risk of cancer in pesticide applicators in Swedish agriculture. *Br J Ind Med* 1989;46:809–14.
46. Rooney C, Beral V, Maconchie N, et al. Case-control study of prostatic cancer in employees of the United Kingdom Atomic Energy Authority. *BMJ* 1993;307:1391–7.
47. Fair WR, Fleshner NE, Heston WY. Cancer of the prostate: a nutritional disease? *Urology* 1997;50:840–8.
48. Corder EH, Friedman GD, Vogelmann JH, et al. Seasonal variation in vitamin D, vitamin D-binding protein, and dehydroepiandrosterone: risk of prostate cancer in black and white men. *Cancer Epidemiol Biomarkers Prev* 1995;4:655–9.
49. Key T. Risk factors for prostate cancer. *Cancer Surv* 1995;23:63–77.
50. Ingles SA, Ross RK, Yu MC, et al. Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. *J Natl Cancer Inst* 1997;89:166–70.
51. Ma J, Stampfer MJ, Gann PH, et al. Vitamin D receptor poly-

- morphisms, circulating vitamin D metabolites, and risk of prostate cancer in United States physicians. *Cancer Epidemiol Biomarkers Prev* 1998;7:385-90.
52. Makela SI, Pylkanen LH, Santti RS, et al. Dietary soybean may be antiestrogenic in male mice. *J Nutr* 1995;125:437-45.
 53. Zhang JX, Hallmans G, Landstrom M, et al. Soy and rye diets inhibit the development of Dunning R3327 prostatic adenocarcinoma in rats. *Cancer Lett* 1997;114:313-14.
 54. Newmark JR, Hardy DO, Tonb DC. Androgen receptor gene mutations in human prostate cancer. *Proc Natl Acad Sci U S A* 1992;89:6319-23.
 55. Taplin ME, Bubley GJ, Shuster TD, et al. Mutation of the androgen-receptor gene in metastatic androgen-independent prostate cancer. *N Engl J Med* 1995;332:1393-8.
 56. Lu J, Danielsen M. A Stu I polymorphism in the human androgen receptor gene (*AR*). *Clin Genet* 1996;49:323-4.
 57. Giovannucci E, Platz EA, Stampfer MJ, et al. The CAG repeat within the androgen receptor gene and benign prostatic hyperplasia. *Urology* 1999;53:121-5.
 58. Mitsumori K, Terai A, Oka H, et al. Androgen receptor CAG repeat length polymorphism in benign prostatic hyperplasia (BPH): correlation with adenoma growth. *Prostate* 1999;41:253-7.
 59. Sawaya ME, Shalita AR. Androgen receptor polymorphisms (CAG repeat lengths) in androgenetic alopecia, hirsutism, and acne. *J Cutan Med Surg* 1998;3:9-15.
 60. Ellis JA, Stebbing M, Harrap SB. Polymorphism of the androgen receptor gene is associated with male pattern baldness. *J Invest Dermatol* 2001;116:452-5.
 61. Mori K, Ushiyama T, Inoue K, et al. Polymorphic CAG repeats of the androgen receptor gene in Japanese male patients with ankylosing spondylitis. *Rheumatology* 2000;39:530-2.
 62. Kooy RF, Reyniers E, Storm K, et al. CAG repeat contraction in the androgen receptor gene in three brothers with mental retardation. *Am J Med Genet* 1999;85:209-13.
 63. Kawasaki T, Ushiyama T, Ueyama H, et al. Polymorphic CAG repeats of the androgen receptor gene and rheumatoid arthritis. *Ann Rheum Dis* 1999;58:500-2.
 64. Yu MW, Cheng SW, Lin MW, et al. Androgen-receptor gene CAG repeats, plasma testosterone levels, and risk of hepatitis B-related hepatocellular carcinoma. *J Natl Cancer Inst* 2000;92:2023-8.
 65. La Spada AR, Wilson EM, Lubahn DB. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 1991;352:77-9.
 66. Kuhlbaumer G, Kress W, Ringelstein EB, et al. Thirty-seven CAG repeats in the androgen receptor gene in two healthy individuals. *J Neurol* 2001;248:23-6.
 67. Yasui T, Akita H, Itoh Y, et al. CAG repeats in the androgen receptor: a case of spinal and bulbar muscular atrophy associated with prostate cancer. *J Urol* 1999;162:495.
 68. Sasaki M, Dahiya R, Fujimoto S, et al. The expansion of the CAG repeat in exon 1 of the human androgen receptor gene is associated modification of *BRCA1*-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat with uterine endometrial carcinoma. *Mol Carcinog* 2000;27:237-44.
 69. Rebbeck TR, Kantoff PW, Krithivas K, et al. *Am J Hum Genet* 1999;64:1371-7.
 70. Spurdle AB, Dite GS, Chen X, et al. Androgen receptor exon 1 CAG repeat length and breast cancer in women before age forty years. *J Natl Cancer Inst* 1999;91:961-6.
 71. Giovannucci E, Stampfer MJ, Krithivas K, et al. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci U S A* 1997;94:3320-3.
 72. Sartor O, Zheng Q, Eastham JA. Androgen receptor gene CAG repeat length varies in a race-specific fashion in men without prostate cancer. *Urology* 1999;53:378-80.
 73. Correa-Cerro L, Wöhr G, Haussler J, et al. CAGnCAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French-German population. *Eur J Hum Genet* 1999;7:357-62.
 74. Edwards SM, Badzioch MD, Minter R, et al. Androgen receptor polymorphisms: association with prostate cancer risk, relapse and overall survival. *Int J Cancer (Pred Oncol)* 1999;84:458-65.
 75. Bratt O, Borg A, Kristoffersson U, et al. CAG repeat length in the androgen receptor gene is related to age at diagnosis of prostate cancer and response to endocrine therapy, but not to prostate cancer risk. *Br J Cancer* 1999;81:672-6.
 76. Eeles RA, Edwards SM, Minter R, et al. Androgen receptor polymorphisms: their association with prostate cancer risk, relapse and overall survival. Programs and abstracts of the 1998 Annual Meeting of the American Society of Human Genetics. *Am J Hum Genet* 1998;4:A21.
 77. Makridakis NM, Reichardt JKV. Molecular epidemiology of hormone-metabolic loci in prostate cancer. *Epidemiol Rev* 2001;23:24-9.
 78. Visakorpi T, Hyytinen E, Koivisto P, et al. In vivo amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* 1995;9:401-6.
 79. Choong CS, Wilson EM. Trinucleotide repeats in the human androgen receptor: a molecular basis for disease. *J Mol Endocrinol* 1998;21:235-57.
 80. Smith PG, Day NE. The design of case-control studies: the influence of confounding and interaction effects. *Int J Epidemiol* 1984;13:356-65.
 81. Cotton SC, Sharp L, Little J, et al. Glutathione S transferase polymorphisms and colorectal cancer. *Am J Epidemiol* 2000;151:846-61.
 82. Bharaj BS, Vassilikos EJ, Diamandis EP. Rapid and accurate determination of (CAG)*n* repeats in the androgen receptor gene using polymerase chain reaction and automated fragment analysis. *Clin Biochem* 1999;32:327-32.
 83. Wada Y, Mitsumori K, Terachi T, et al. Measurement of polymorphic trinucleotide repeats in the androgen receptor gene by matrix assisted laser desorption/ionization time-of-flight mass spectrometry. *J Mass Spectrom* 1999;34:885-8.
 84. Lange EM, Chen H, Brierley K, et al. The polymorphic exon 1 androgen receptor CAG repeat in men with a potential inherited predisposition to prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2000;9:439-42.

APPENDIX. INTERNET SITES

Prostate cancer

American Cancer Society:

<http://www.cancer.org>

CaP Cure:

<http://www.capcure.org>

National Cancer Institute:

<http://www.cancer.gov>

National Prostate Cancer Coalition:

<http://www.4npcc.org>

University of Pennsylvania Cancer Center OncoLink:

<http://oncolink.upenn.edu/disease/prostate>

Genetic databases

The Androgen Receptor Gene Mutations Database:

<http://www.mcgill.ca/androgendb>

Online Mendelian Inheritance in Man (OMIM):

<http://www.ncbi.nlm.nih.gov/Omim> (for PCA [MIM 176807], AR [MIM 313700], *HPC1* [MIM 601518], *HPC2* [MIM 602759], and *HPCX* [MIM 300147])

The Genome Database (GDB)

<http://gdbwww.gdb.org>

<http://gdb.org/gdb-bin/eneral/accco?accessionNum-GDB:120556> (AR)